

Micromagnetic simulations on GPU, a case study: vortex core switching by high-frequency magnetic fields

B. Van de Wiele¹, A. Vansteenkiste², M. Kammerer³, B. Van Waeyenberge²,
L. Dupré¹, and D. De Zutter⁴

¹ Dept. of Electrical Energy, Systems and Automation, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium.

² Dept. of Solid State Physics, Ghent University, Krijgslaan 281-S1, 9000 Gent, Belgium.

³ Max Planck Institute for Intelligent Systems, Heisenbergstrasse 3, 70569, Stuttgart, Germany.

⁴ Dept. of Information Technology, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium.
ben.vandewiele@ugent.be

Abstract—Since magnetic vortex cores have two ground states, they are candidates for digital memory bits in future MRAM devices. Vortex core switching can be induced by exciting the gyrotropic eigenmode e.g. by applying cyclic magnetic fields with typically a sub-GHz frequency. However, recent studies reveal that other modes exist that can be excited at higher frequencies, but still lead to switching with relatively small field amplitudes. Here, we perform a full scan of the frequency/amplitude parameter space to explore such excitation modes. The enormous amount of simulations can only be performed in an acceptable time span when the micromagnetic (CPU-) simulations are drastically accelerated. To this aim, we developed MUMAX, a GPU based software tool that speeds up micromagnetic simulations with about two orders of magnitude compared to standard CPU micromagnetic tools. By exploiting MUMAX’ numerical power we were able to explore new switching opportunities at moderate field amplitudes in the frequency range between 5 and 12 GHz.

I. INTRODUCTION

Magnetic vortices occur in micron and submicron sized thin-film soft magnetic structures of suitable size. The vortex has a flux-closed magnetization state, making it energetically stable. At the center on the other hand, its magnetization is forced to be perpendicular to the plane in order to minimize the exchange energy. The so-formed *vortex core* is only a few exchange lengths in diameter —typically tens of nanometers [1]— and has two ground states denoted by *up* and *down*. Since the discovery of vortex core switching by excitation of its so-called gyrotropic mode [2], interest has arisen in using vortices as magnetic memory bits for future MRAM applications [3].

The original switching method [2] consisted of applying a magnetic field oscillating along one in-plane axis. Here, the frequency is tuned to the gyrotropic eigenmode (typically in the 100’s of MHz-range [4]). This resonant excitation is thus relatively slow. Currently, other excitation schemes are being explored as well. E.g., excitation with rotating magnetic fields provide full control over the core polarization [5], while excitation with short magnetic field pulses is well suited for fast switching [6]. Switching by using electrical currents exploiting the spin-transfer torque effect is also possible [7], [8]. Recently, interest has arisen in switching with high-frequency (several GHz) rotating fields [9]–[11]. These fields excite vortex modes with higher frequencies than the gyrotropic mode and may provide faster switching mechanisms. In all of these works micromagnetic simulations have been successfully

employed to predict or explain the vortex core switching behavior.

When simulating this kind of switching, a large parameter space manifests itself: both the amplitude and the frequency (or pulse length) of the oscillating field can be varied, but also different nanostructure sizes and material parameters might be considered. Hence, this kind of research needs very time efficient micromagnetic simulators. In this context, CPU based computations often limit the research possibilities due to long simulation times. In this contribution, we simulate magnetic vortex core switching by magnetic fields oscillating along one in-plane axis in the GHz regime with the new MUMAX micromagnetic simulator running on Graphics Processing Units (GPUs) [12]. This offers a speedup of up to a factor of 100 compared to, e.g. the CPU based OOMMF software package [13]. This speedup allows to fully simulate very large parameter spaces by brute force in search of interesting physical phenomena.

II. MUMAX: GPU PARALLELIZED MICROMAGNETICS

When developing micromagnetic simulation software, time and memory efficiency —together with accuracy— are key concerns. Indeed, due to the very small space and time discretizations imposed by the theory, micromagnetic simulations result typically in very large numerical problems. Therefore efficient algorithms have been introduced to alleviate the computational bottle necks: the evaluation of the demagnetizing field has been accelerated resulting in $\mathcal{O}(N \log N)$ [14], [15] and $\mathcal{O}(N)$ [16] algorithms instead of $\mathcal{O}(N^2)$ where N is the number of discretization cells. Furthermore time efficient time stepping schemes have been developed to reduce the number of field evaluations. While nowadays no large speedups are to be expected any more from the development of new algorithms, efficiency gains are still to be achieved by exploiting the fast hardware evolutions.

Parallel computing has shown to be a very good solution for accelerating numerical software. Here, the time gains are determined by the degree to which the code can be parallelized, the number of concurrent threads and the communication cost. In general, micromagnetic codes can be easily partitioned in portions which can run in parallel. Multithreaded micromagnetic software distributes the work load over multiple cores present in one machine [17]. Here, the number of parallel

threads (i.e. the number of cores) is relatively small, but the communication is fast. MPI-based software distributes the work load over various machines present in a cluster [18]. In the cluster, the number of concurrent threads can be very large, but the parallel efficiency strongly depends on the used communication network. Speedups are only satisfactory when an infiniband network is available which results in a high financial cost.

In the last years, GPUs have gained popularity in the numerical programming community. Although originally designed for purely graphical purposes, GPUs turn out to be extremely well-suited for high-performance calculations. The GPU hardware consists of several multiprocessors which contain a large number of computational cores. This way, an enormous amount of $\mathcal{O}(10^4)$ of threads can run parallel jobs. Furthermore, since the multiprocessors and the GPU memory are all placed on the same device, the data communication within the GPU is extremely fast with a typical bandwidth of a few hundred GB/s. This, combined with the fact that very large portions of the micromagnetic code can be parallelized make a GPU based parallelization of micromagnetic code extremely advantageous.

To exploit the huge parallel power of GPUs, the numerical algorithms and their implementation should be completely reconsidered in order to fit the characteristic GPU architecture. Since data transfer between the CPU memory and GPU is about two orders of magnitude slower than the communication between the GPU and its own dedicated memory (called global memory), all computations should be performed by the GPU. A sketch of the work flow in a general GPU based code is given in Fig. 1. When implementing the GPU jobs one needs to take into account several precautions to optimize the GPU performance. As a main concern, all threads within a warp (a group of threads) should perform exactly the same computations (on different data) and all the read and write instructions are to be performed in a coalesced way, i.e. the threads within a warp should access corresponding data elements in prescribed blocks of the GPU memory. Only then, the full parallel computational force is used. During the development of the GPU based micromagnetic simulator MUMAX, these implementation restrictions are carefully taken into account.

The current version of MUMAX parallelizes only over the cores of a single GPU card, not over multiple GPUs or multiple networked machines. So in principle, a running simulation requires very few data transfers to and from the GPU. In our implementation, the simulation starts by sending the initial magnetization state from the CPU to the GPU. A main loop, running on the CPU, instructs the GPU step-by-step what to do with this data. Each such step is a GPU subroutine, implemented in C for CUDA. The GPU acts only on (large) data arrays in its dedicated memory. Only when data needs to be output to disk, large data transfers from GPU to CPU inevitably take place. Operations on small amounts of numbers, which can not be massively parallelized, still happen on the CPU. E.g., after having taken a time step to solve the Landau-Lifshitz equation the CPU calculates the optimal time step for the next iteration and passes this single number back

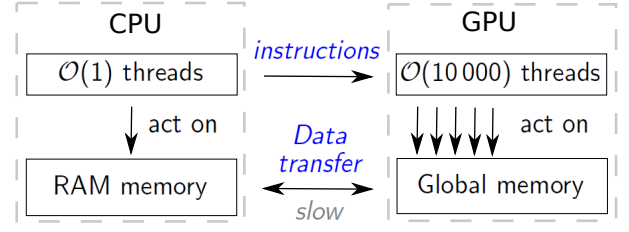


Fig. 1. Sketch of the workflow for GPU based computations. On a multicore CPU a few threads can work in parallel, while on a GPU tens of thousands of parallel threads are available. A program always starts on the CPU and sends instructions to the GPU. Usually CPU and GPU have separate global RAM memories. Both also have fast, but scarce, on-chip memory: the CPU cache memory and GPU shared memory, respectively. Compared to the internal GPU communication, the data transfer between the CPU and the GPU is about two orders of magnitude slower. Hence, in an efficient GPU code the CPU-GPU data transfer should be minimized.

to the GPU. In this way, much of the code can still run on the CPU without notable communication overhead.

In MUMAX, a finite difference space discretization is applied. The convolution product defined by the magnetostatic field is accelerated using fast Fourier transforms (FFTs). However, the GPU implementation of these FFTs does not use the standard real to complex transforms available in the CUFFT library. Instead, a dedicated FFT routine is developed in which, e.g., the 3D FFTs are broken up in sets of 1D transforms in the three dimensions. In this way, one can avoid 1D FFTs on arrays that contain only zeros originating from zero padding. Furthermore, it is most advantageous to perform the 1D FFTs on contiguous data arrays. Therefore, efficient matrix transposes are required that make use of so called *shared memory* to ensure coalesced reading and writing. This FFT scheme is much faster than the built-in 3D CUFFT routine.

Since MUMAX is designed to be an easily accessible open source micromagnetic simulation tool¹, free to use by the magnetic community, the code can also run in CPU mode. On one core, the efficiency is comparable to the OOMMF software. However, in order to also fully exploit the computational power on CPU, the code can run in parallel on several CPU cores by means of multithreaded commands. The performance of both parallel implementations –GPU and multithreaded CPU– is shown in Fig. 2. Here, the speedups are shown as obtained in GPU-mode and in CPU-mode on 4 cores compared to single core CPU computations. It is clear that the speedups obtained with the GPU implementation are enormous. In particular speedups for large dimensions corresponding with a power of 2 exceed a factor 100. These sizes perfectly fit the GPU architecture as well as the nature of the FFTs. Although completely different, also the speedups obtained with the multithreaded CPU implementation are close to optimal: up to a factor 3.5 when 4 cores are used. It is clear that the massive time gains obtained with MUMAX enable a thorough numerical investigation of large parameter spaces e.g. the RF switching of vortices considered here. Furthermore, it should be noted that the software is validated

¹MUMAX can be downloaded from <http://dynamat.ugent.be/mumax>

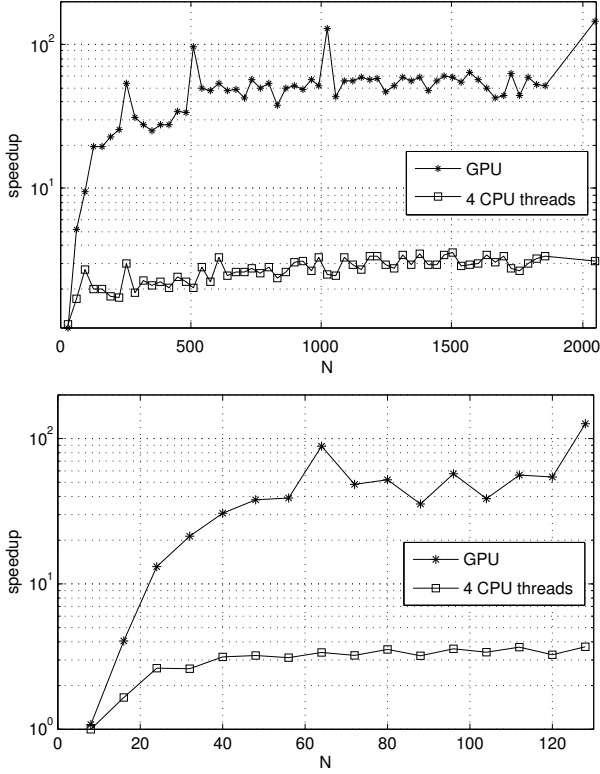


Fig. 2. Speedup obtained for 2D geometries with dimensions $N \times N$ (top) and 3D geometries with dimensions $N \times N \times N$ (bottom). The GPU computations are performed on a GTX580 GPU, while the CPU computations are performed on a 2.8GHz core i7-930 CPU.

by comparing simulation results for several micromagnetic standard problems [12].

III. SIMULATIONS RESULTS

We simulate the vortex dynamics in a 500 nm wide, 50 nm thick Permalloy ($\text{Ni}_{80}\text{Fe}_{20}$) disk with a saturation magnetization of 800 kA/m, zero anisotropy and damping coefficient equal to 0.01. Discretization cells with dimensions $3.90625 \text{ nm} \times 3.90625 \text{ nm}$ are used. To this sample, a broad range of magnetic fields oscillating along the horizontal axis with varying frequencies and amplitudes is applied. Since we are interested in fast switching methods, we apply only 10 oscillation periods: if the vortex core has not yet switched during those first 10 periods, the mechanism is rather slow and not very interesting for applications.

We investigate the spatially averaged in-plane magnetization of the vortex. This quantity is correlated with the displacement of the vortex core from the structure center and thus serves as a measure for the *response* of the system to an excitation [19]. Figure 3 shows the simulated response as a function of excitation frequency at a fixed low field amplitude of 1 mT, which is insufficient to cause switching. The response shown in Fig. 3 clearly exhibits a number of resonance peaks. The lowest-frequency peak (700 MHz) corresponds to the vortex gyrotropic mode, while at higher frequencies (7.3 GHz, 10.5 GHz, 12.2 GHz) additional resonances appear. These correspond to azimuthal spin waves with different numbers of

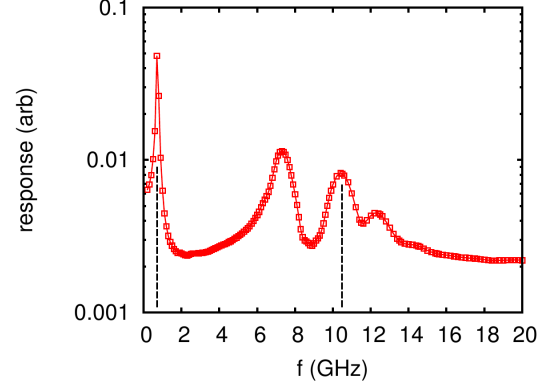


Fig. 3. Response (magnitude of the spatially averaged in-plane magnetization) of a vortex in a 500 nm wide, 50 nm thick Permalloy disk, when excited by a low-amplitude oscillating magnetic field of variable frequency. Various resonances can be observed. The lowest-frequency resonance of 700 MHz corresponds to the vortex gyrotropic mode, and higher resonances are present at 7.3 GHz, 10.5 GHz and 12.2 GHz. The dashed lines correspond to those in Fig. 4.

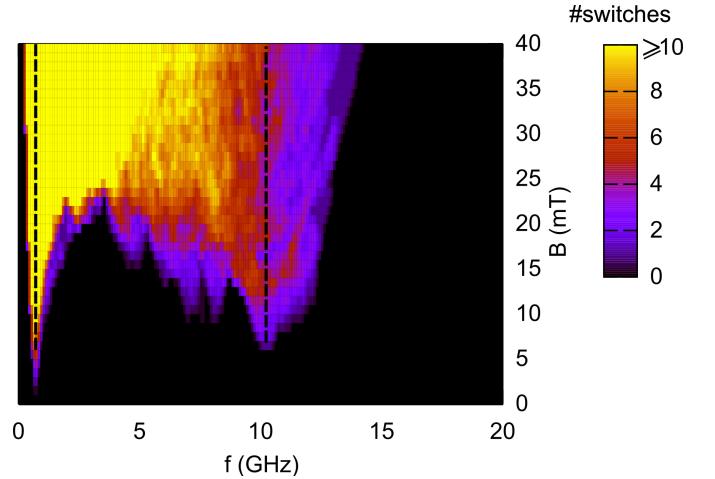


Fig. 4. Vortex core switching behavior in a 500 nm wide, 50 nm thick Permalloy disk when excited by 10 periods of a magnetic field oscillating along the horizontal axis with variable frequency and amplitude. The color corresponds with the number of switches (black means no switching). At a few frequencies an upside-down *peak* occurs where the required amplitude for switching reaches a minimum. At 700 MHz this corresponds to the gyrotropic mode. The peak at 10.5 GHz is most interesting as it corresponds to a high-frequency excitation for which a relatively low field of only about 7 mT is sufficient for switching. The dashed lines correspond to those in Fig. 3.

nodes, travelling around the vortex core (see, e.g., [20]). It is well known that the lowest (gyrotropic) mode can lead to vortex core switching when sufficiently excited [2]. However, the excitation modes in the GHz-range are only recently used – by applying *rotating* magnetic fields – to switch the vortex core [11]. Now we elaborate a full scan of the frequency/amplitude parameter space of magnetic fields oscillating *along a fixed axis*, in order to excite the above mentioned modes and investigate their ability to initiate switching.

The amplitude and frequency ranges considered here are 0-40 mT (1 mT steps) and 0 - 20 GHz (100 MHz steps) respectively. Figure 4 shows the number of vortex core switches as a function of the excitation parameters. This is mainly

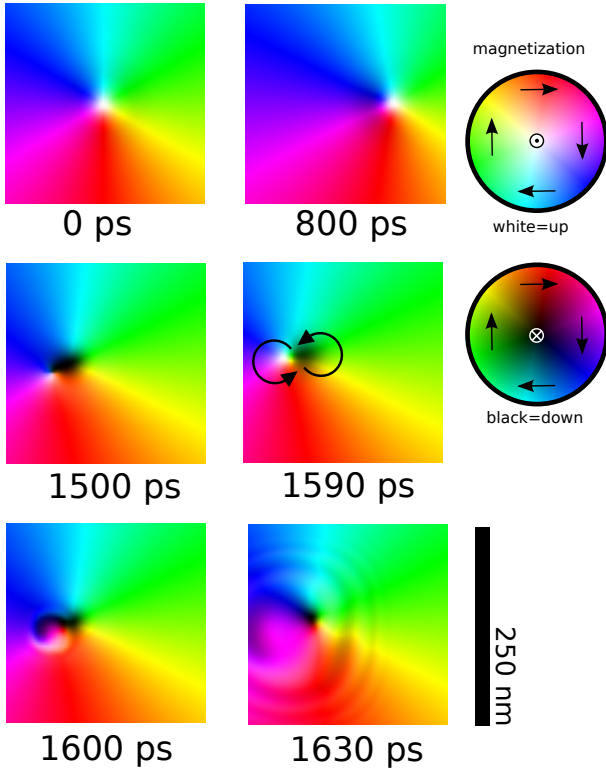


Fig. 5. Representative magnetization snapshots of the center part of a 500 nm wide, 50 nm thick Permalloy disk excited by a 700 MHz, 5 mT oscillating field. In the top row, the system is gradually absorbing energy from the field resulting in a large gyrotropic movement of the vortex core (white spot). An oppositely magnetized region (*dip*, black spot) starts to form next to the vortex core. The dip evolves into a vortex-antivortex pair leading to the annihilation of the original up vortex and leaving a down vortex (black dot) behind. The annihilation process releases a burst of short-wavelength spin waves, visible in the last frame. The curling arrows indicate the vortices (original and new). The antivortex is located at the intersection of the arrows.

intended to illustrate whether or not the vortex core switches for a certain frequency/amplitude combination. In the entire black region, no switching occurs. The most prominent region where switching does occur is the leftmost peak around 700 MHz, corresponding to the gyrotropic mode. When going from the gyrotropic region to higher frequencies, increasingly higher amplitudes are necessary to achieve switching. Indeed, this corresponds to de-tuning the excitation from the gyrotropic resonance to regions where the system responds less to the field. However, between 5 and 12 GHz, the threshold for switching starts to decrease again. This region corresponds to the three broad resonances visible in Fig. 3. As the vortex responds again more strongly at these frequencies, the corresponding switching threshold decreases again. Most notably, at 10.5 GHz a relative minimum of only about 7 mT is reached. Hence, from an energetic point of view this region is very interesting since it combines relatively low field amplitudes with high-frequency —and thus fast— switching behavior.

Figures 5 and 6 show more details of the dynamics in the gyrotropic and the high-frequency region, respectively.

The gyrotropic switching illustrated in Fig. 5 is already well-known. By excitation, the vortex starts to gyrate around its equilibrium position, which causes a dynamic deformation

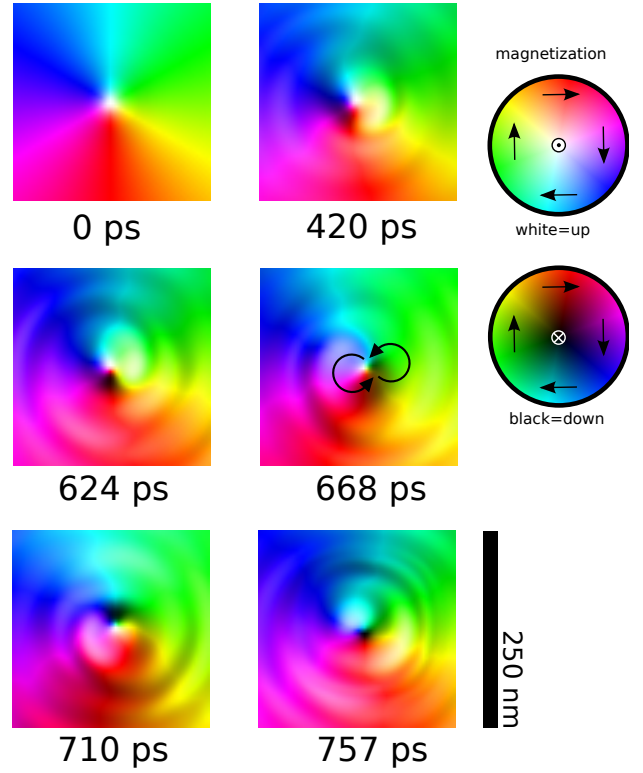


Fig. 6. Representative magnetization snapshots of the center part of a 500 nm wide, 50 nm thick Permalloy disk excited by a 12 GHz, 10 mT oscillating field. In the top row, the system is gradually absorbing energy from the field resulting in a small movement of the vortex core (white spot) and a prominent emission of spin waves. An oppositely magnetized region (*dip*, black spot) starts to form next to the vortex core. The dip evolves into a vortex-antivortex pair leading to the annihilation of the original up vortex and leaving a down vortex behind. The annihilation process releases a burst of short-wavelength spin waves, visible in the last frame. The curling arrows indicate the vortices (original and new). The antivortex is located at the intersection of the arrows.

of the core. This so-called *dip* has a polarization opposite to the original core [21]. As soon as the movement reaches a certain critical velocity [7], a vortex-antivortex pair is created which leads to vortex core switching [2].

The difference between the well-known gyrotropic switching dynamics and those in the GHz regime is highlighted by Fig. 6. Here, a 12 GHz field is applied. It can be seen that the movement of the vortex core is only minuscule. Contrary to the gyrotropic case, spin waves are now radiated strongly from the vortex core. It can be seen that despite the absence of a high vortex core velocity, a *dip* is nevertheless formed. As soon as the dip is strong enough, switching through the same vortex-antivortex mechanism happens as in the gyrotropic case. Similar spin wave induced switching behavior was reported and explained for excitation with rotating magnetic fields [11].

Because of the elaborate parameter scan, we were able to identify vortex core switching events in a broad frequency range. This was only made possible thanks to the GPU-based micromagnetic simulations. In total, 8000 separate simulations were performed on a single machine containing four nVIDIA

GTX 480 GPUs². This huge batch of simulations completed in only 20 hours, which is remarkably fast considering the fine parameter scan. It is estimated that the same simulations would have taken on the order of a month on a regular CPU.

IV. CONCLUSION

We have simulated vortex core switching in Permalloy nanodisks excited by magnetic fields oscillating along one axis with frequencies up to the several GHz range. Apart from switching at the gyrotropic eigenfrequency—which is well-known—we also investigated high-frequency resonances at which switching is possible with relatively low magnetic field amplitudes. These switching paths might be of interest for future vortex-MRAM applications where switching speeds need to be as fast as possible with minimal excitation field amplitudes. We could obtain these results thanks to a detailed scanning of the large frequency/amplitude parameter space. Such elaborate simulations are made possible by graphical processing hardware that allow to speed up simulations by a factor of 100 compared to regular CPU based programs. This paper demonstrates that fast GPU-based micromagnetic simulation tools as MUMAX can contribute significantly to the design of future magnetic applications.

ACKNOWLEDGMENTS

This research is financially supported by the Flanders Research Foundation (FWO) and also by FWO project G.0142.08.

REFERENCES

- [1] K. W. Chou, A. Puzic, H. Stoll, D. Dolgos, G. Schütz, B. Van Waeyenberge, A. Vansteenkiste, T. Tyliczszak, G. Woltersdorf, and C. H. Back. Direct observation of the vortex core magnetization and its dynamics. *Appl. Phys. Lett.*, 90:202505, 2007.
- [2] B. Van Waeyenberge, A. Puzic, H. Stoll, K. W. Chou, T. Tyliczszak, R. Hertel, M. Fähnle, H. Bruckl, K. Rott, G. Reiss, I. Neudecker, D. Weiss, C. H. Back, and G. Schütz. Magnetic vortex core reversal by excitation with short bursts of an alternating field. *Nature*, 444:461–464, 2006.
- [3] S. Bohlens, B. Kruger, A. Drews, M. Bolte, G. Meier, and D. Pfannkuche. Current controlled random-access memory based on magnetic vortex handedness. *Applied Physics Letters*, 93(14):142508, October 2008.
- [4] B. E. Argyle, E. Terrenzio, and J. C. Slonczewski. Magnetic vortex dynamics using the optical Cotton-Mouton effect. *1984 Digests Of Intermag '84. International Magnetism Conference (cat. No. 84ch1918-2)*, pages 350–350, 1984.
- [5] M. Curcic, B. Van Waeyenberge, A. Vansteenkiste, M. Weigand, V. Sackmann, H. Stoll, H. Fähnle, T. Tyliczszak, G. Woltersdorf, C. H. Back, and G. Schütz. Polarization selective magnetic vortex dynamics and core reversal in rotating magnetic fields. *Physical Review Letters*, 101(19):197204, 2008.
- [6] M. Weigand, B. Van Waeyenberge, A. Vansteenkiste, M. Curcic, V. Sackmann, H. Stoll, T. Tyliczszak, K. Kaznatcheev, D. Bertwistle, G. Woltersdorf, C. H. Back, and G. Schütz. Vortex core switching by coherent excitation with single in-plane magnetic field pulses. *Physical Review Letters*, 102(7):077201, February 2009.
- [7] K. Yamada, S. Kasai, Y. Nakatani, K. Kobayashi, H. Kohno, A. Thiaville, and T. Ono. Electrical switching of the vortex core in a magnetic disk. *Nat. Mater.*, 6:269–273, 2007.
- [8] E. Chen, D. Apalkov, Z. Diao, A. Driskill-Smith, D. Druist, D. Lottis, V. Nikitin, X. Tang, S. Watts, S. Wang, S. A. Wolf, A. W. Ghosh, J. W. Lu, S. J. Poon, M. Stan, W. H. Butler, S. Gupta, C. K. A. Mewes, T. Mewes, and P. B. Visscher. Advances and future prospects of spin-transfer torque random access memory. *IEEE Transactions On Magnetics*, 46(6):1873–1878, 2010.
- [9] V. P. Kravchuk, D. D. Sheka, Y. Gaididei, and F. G. Mertens. Controlled vortex core switching in a magnetic nanodisk by a rotating field. *Journal of Applied Physics*, 102(4), 2007.
- [10] V. P. Kravchuk, Yu. Gaididei, and D. D. Sheka. Nucleation of a vortex-antivortex pair in the presence of an immobile magnetic vortex. *Physical Review B (Condensed Matter and Materials Physics)*, page 100405 (4 pp.), 2009.
- [11] M. Kammerer, M. Weigand, M. Curcic, M. Noske, M. Sproll, A. Vansteenkiste, B. Van Waeyenberge, H. Stoll, G. Woltersdorf, C. H. Back, and G. Schuetz. Magnetic vortex core reversal by excitation of spin waves. *Nat Commun*, 2:279–, April 2011.
- [12] A. Vansteenkiste and B. Van de Wiele. Mumax: a new high-performance micromagnetic simulation tool. *Submitted to: Journal of Magnetism and Magnetic Materials*, 2011.
- [13] MJ Donahue and DG Porter. OOMMF user's guide, version 1.0. interagency report NISTIR 6376, national institute of standards and technology, gaithersburg, MD, 1999.
- [14] B. Van de Wiele, F. Olyslager, L. Dupre, and D. De Zutter. On the accuracy of fft based magnetostatic field evaluation schemes in micromagnetic hysteresis modeling. *Journal of Magnetism and Magnetic Materials*, pages 469–76, 2010.
- [15] H. H. Long, E. T. Ong, Z. J. Liu, and E. P. Li. Fast fourier transform on multipoles for rapid calculation of magnetostatic fields. *IEEE Transactions On Magnetics*, 42(2):295–300, 2006.
- [16] G. Brown, T. C. Schulthess, D. M. Apalkov, and P. B. Visscher. Flexible fast multipole method for magnetic simulations. *IEEE Transactions On Magnetics*, 40(4):2146–2148, 2004.
- [17] M. J. Donahue. Parallelizing a micromagnetic program for use on multiprocessor shared memory computers. *IEEE Transactions On Magnetics*, 45(10):3923–3925, 2009.
- [18] W. Scholz, J. Fidler, T. Schrefl, D. Suess, R. Dittrich, H. Forster, and V. Tsiantos. Scalable parallel micromagnetic solvers for magnetic nanostructures. *Computational Materials Science*, 28(2):366–383, 2003.
- [19] M. Najafi, B. Kruger, S. Bohlens, M. Franchin, H. Fangohr, A. Vanhaverbeke, R. Allenspach, M. Bolte, U. Merkt, D. Pfannkuche, D. P. F. Moller, and G. Meier. Proposal for a standard problem for micromagnetic simulations including spin-transfer torque. *Journal of Applied Physics*, 105(11), 2009.
- [20] M. Buess, R. Hollinger, T. Haug, K. Perzlmaier, U. Krey, D. Pescia, M. R. Scheinfein, D. Weiss, and C. H. Back. Fourier transform imaging of spin vortex eigenmodes (vol 93, art no 077207, 2004). *Phys. Rev. Lett.*, 93:129902, 2004.
- [21] A. Vansteenkiste, K.W. Chou, M. Weigand, M. Curcic, V. Sackmann, H. Stoll, T. Tyliczszak, G. Woltersdorf, C.H. Back, G. Schütz, and B. Van Waeyenberge. X-ray imaging of the dynamic magnetic vortex core deformation. *Nature Physics*, 5:332–334, 2009.

²The used GTX 480 GPUs each have 480 CUDA cores clocked at 1.4 GHz. They are grouped in 15 multiprocessors of 32 cores. The 1.5 GB global memory is clocked at 1.8 GHz and has a 384 bit wide interface delivering a bandwidth of 177 GB/s.